

BOOK REVIEW: MCINTYRE, D. & WALKER, B. (2019) *CORPUS STYLISTICS: THEORY AND PRACTICE*. EDINBURGH UNIVERSITY PRESS.

SALLY HUNT
OPEN UNIVERSITY

CITATION

Hunt, S. (2021). Book Review: McIntyre,
D. & Walker, B. (2019) *Corpus Stylistics:
Theory and Practice*. Edinburgh University
Press.. Journal of Corpora and Discourse
Studies, 4(1):63–67

CONTACT

Sally Hunt, School of Languages and Applied Linguistics, Open University, UK.
Sally.Hunt@open.ac.uk

DOI

<https://doi.org/10.18573/jcads.69>

ORCID

0000-0001-5800-8118

ISSUE DOI

<https://doi.org/10.18573/jcads.v4>

LICENSE

© The authors. Available under the terms of the CC-BY 4.0 license

Review accepted 2021-05-06

Book Review: McIntyre, D. & Walker, B.
(2019) *Corpus Stylistics: Theory and Practice.*
Edinburgh University Press

Sally Hunt
Open University

This timely work in a relatively new field should be useful to researchers experienced in both stylistics and corpus linguistics, despite being explicitly aimed at stylisticians. In addition to demonstrating the usefulness of incorporating corpus methods into stylistic enquiry, there is much to convince corpus linguists that applying their usual toolkit to stylistics is itself a worthwhile pursuit. The book would be worth recommending to students as a helpful introduction to the area of corpus stylistics.

In delineating its scope, *Corpus Stylistics* provides clear introductions to key concepts and methodological processes, especially those in corpus linguistics. The book begins by asking ‘what is corpus stylistics?’, and returns regularly to questions of definition and delineation in subsequent chapters, while demonstrating applications with both fiction and non-fiction texts. In the final chapter, the authors revisit the problem of definition, and conclude that corpus stylistics best describes “a particular focus of corpus linguistics” which is concerned with “issues of style” (p. 315). That corpus stylistics is described as a focus of corpus linguistics, rather than as one way of *doing* stylistics, did surprise me somewhat, as it might alienate some stylisticians. Perhaps the stylistics ‘credentials’ of the authors, and the way this definition was carefully woven from previous theory and analysis, might soften the blow and justify the apparent primacy given to corpus linguistics. This ‘bookend’ structure neatly frames the book and its central concerns with a depth and scope which reflect the authors’ expertise, and helps to establish corpus stylistics as a field separate from related applications of corpus linguistics, such as corpus-assisted discourse studies (CADS), for instance.

The structure of the chapters makes the book especially useful and helps to prevent overwhelming the newcomer to corpus methods. Important theoretical and methodological concepts are sprinkled across chapters as appropriate, thus avoiding beginning the book with a concentration of material new to stylisticians, which might potentially dishearten the reader. So, for example, introductions to best practice in building and annotating corpora, sampling and ethics, the concepts of keyness and collocation, and the use of corpora to test hypotheses, are dealt with in chapters 2 to 4, as they arise. Later chapters focus on specific contexts or text types of stylistic interest, e.g. classroom or historical texts. The case studies found in most chapters have been well chosen to illustrate the specific aspect in focus or to concretise the theoretical or methodological choices. One might expect these case studies to be based on fictional texts, and they frequently are, but

not exclusively, and they reflect the authors' stance that stylistics is not only concerned with canonical literature (see, for example, p. 309ff), since they also analyse a number of political texts and 2000AD's Judge Dredd comic strips.

Chapter 1 is, by necessity, introductory, with an explanation of what is meant by a corpus and a sketch of the development of corpus linguistics, which, although brief, is neither thin nor lacking in depth, and covers the major contributions to the field. This chapter also begins the thread which runs through the book: the scope of corpus stylistics. This is introduced via a discussion of the emergence of the area, and through contextualising it with reference to neighbouring approaches: stylistics itself, and cognitive stylistics. The authors show how parallel developments in stylistics and in the more Firthian corpus approaches have blended in corpus stylistics, offering researchers possibilities sometimes previously thought impossible in the field of stylistics, such as the calculation of 'background norms', in literary texts especially, against which the features of a particular text can be explored.

Freeman's 1970 statement that, even if it were possible to ascertain these background norms in terms of frequencies and transitional probabilities, they would offer "no particularly revealing insight into either natural language or style" (1970: 6, cited in Walker and McIntyre 2019, p. 26), acts as the springboard for Chapter 2. The authors aim to show that this is not true, while continuing their exploration of the terms used to label the area, and the concepts within it, such as 'corpus-assisted analysis', which they rename 'corpus-informed stylistics'. The case studies show how qualitative stylistics can be augmented with corpus techniques and concepts to show patterns of stylistic interest, and include examples of words and phrases used in meaningfully patterned ways, often in large general corpora of natural language. Le Carré's novel collocation "strolling with intent" is interpreted with reference to the general usage of 'with intent' in the British National Corpus (BNC) (Davies 2004), for instance, and the reader is encouraged to walk through the analysis by following a step-by-step guide to accessing the BNC online. More complex searches are similarly broken down into simple steps and their stylistic implications considered, providing the newcomer to corpus methods with a gentle, but inspiring, introduction to the approach, which provides ample evidence to counter Freeman's claim.

Chapter 3 is concerned with best practice in the construction and annotation of specialised corpora, contrasting the use of ready-made corpora for what the authors term corpus-informed (or corpus-assisted) work (as explored in Chapter 2) to a corpus-based approach, in which the corpus is designed by the researcher for direct analysis. The difference between the two is important for stylistics, because while corpus-informed analysis allows one to check general usage in order to contextualise the particular use of a phrase or word in the (fictional) text in focus, building a corpus of a particular author's work, for instance, allows for direct analysis of stylistic patterns. Representativeness, sampling and balance are considered here with reference to corpus construction, and the

implications of the researcher's decisions drawn out, such as the choice between whole text collection (following Sinclair) and sampling a set number of words from each of a collection of texts (as used in the Lancaster-Oslo-Bergen [LOB] family of corpora of varieties of English (Johansson, Leech and Goodluck 1978)). Copyright and ethics are discussed as concepts central to the process of building corpora.

Annotation – the addition of descriptive labels and classification to corpus data – is also explored here in some depth, with a useful graphic showing the kinds of meta-data and where they would be found. The addition of annotation can be especially revealing of patterns of the feature highlighted. The case study in this chapter is a diachronic comparison of the 2000AD cartoon 'Judge Dredd', with Extensible Mark-up Language (XML) where mark-up is added to identify textual categories particular to comics, such as thought balloons versus speech balloons, and captions and sound effects. The frequencies of each are normalised to enable comparison between comic strips written 25 years apart. As a discourse analyst, I would have been interested in exploring the lexical frequencies in more detail, and the collocations of some central words, but in terms of illustrating the usefulness of annotation, the case study went far enough.

The distinction between research driven by hypotheses and 'unfocused looking' when working with a corpus is returned to in Chapter 4. Here the authors reconsider the terms corpus-based and corpus-driven: how corpora may be used to test hypotheses or answer questions using existing theories, or explored in their raw form, for patterns emerging from the data, unconstrained by theoretical assumptions, such as the categorisation implicit in Part of Speech (PoS) tagging, for instance. The discussion of this distinction, and its implications, is both satisfyingly nuanced and concrete, and is based on studies which will be familiar to many corpus linguists. The role of intuition, in both stylistics and corpus linguistics, is recognised. The rest of the chapter shows how to generate research questions and hypotheses in corpus stylistics, and so improve rigour and objectivity in stylistics through the addition of corpus methods. Useful tools from Wordsmith V7 (Scott 2016) and W-Matrix (Rayson 2009) are used to answer research questions with regard to the work of Dickens and Hemingway. For example, the Standardised Type-Token Ratio in Wordsmith Tools 7 is used to answer questions about the size of Dickens' vocabulary.

Statistics is the focus in Chapter 5, ranging from descriptive measures, like frequency, to inferential statistics, such as keyness and measures of collocational strength. These are complex concepts and those with a traditional literary background may find these calculations and what they mean a little foreign. However, how they work and their implications for stylistic analysis are very clearly and practically explained, as is characteristic of this book. As in other chapters, worked examples walk the reader through the analyses step-by-step, and the case study demonstrates the use of keyness as a rationale for selecting portions of a longer text, in this case David Peace's book *1974*, for qualitative analysis.

Chapter 6 represents a shift in focus which continues for four chapters: to address central concerns in stylistics and how they may be investigated using corpus methods. In this vein, the current chapter considers characterisation in intralingual subtitling, with the novel *Tinker Tailor Soldier Spy* forming the subject of the case study. I particularly enjoyed the inclusion of insights from conversation analysis and pragmatics in the analysis of the discourse marker ‘oh’.

The inclusion of corpus methods in pedagogical stylistics is the topic of Chapter 7, which considers what stylistics offers pedagogically in first-, second-, and foreign-language teaching classrooms (as opposed to the pedagogy of stylistics) and then how language teaching may incorporate corpus-based activities. The Adrian Henri poem *Tonight at Noon* is analysed in some detail, using corpus-informed stylistics, to demonstrate the exploration of ‘language norms in Standard English’ (McIntyre and Walker 2019, p. 209), including the use of collocates, concordance lines and reference corpora.

Historical corpus stylistics is covered in Chapter 8, which also illustrates concepts introduced in earlier chapters, such as corpus construction and annotation, while exploring diachronic shifts in aspects of particular interest to stylistics, such as the representation of speech, writing and thought (also known as discourse presentation, or DP). There are two case studies in this chapter: the first considers DP in Early Modern English (EME) texts versus a corpus of twentieth century fiction, while the second concerns modal auxiliary verbs in EME journalism. To the researcher not trained in stylistics, or, in my case, a corpus methods devotee who finds literary and historical texts interesting from a Critical Discourse Analysis point of view, these last two chapters proved enticing, suggesting socially interesting aspects of language use which I had not previously considered investigating via corpus methods.

While acknowledging the problematic practice of measuring the value of research by its impact, the authors focus in Chapter 9 on ‘real-world’ applications for corpus stylistics, and on non-literary texts. In addition to the implications the reader may draw from this, the use of corpus tools, in particular W-Matrix (Rayson 2009) and AntConc (Anthony 2018), to reveal patterns in political texts is as revealing of the representation of human concerns and relations as the analysis of any literary text. The authors guide readers through the use of potentially bewildering tables of semantic categories and keywords as smoothly and clearly as in earlier chapters, pointing out the significant implications they find, and produce an interpretation which is thoroughly supported by textual evidence and statistical measures. This is, after all, a major strength of the use of corpus tools for any interpretative analysis: being able to substantiate one’s conclusions with evidence, and this chapter should leave any wavering stylistician in no doubt as to its usefulness.

The book began by acknowledging the difficulties inherent in a new area of scholarly inquiry: the “lack of a clear definition [of corpus stylistics] and its as yet not fully developed analytical methods” (McIntyre and Walker 2019, p. 1) and set about addressing

these comprehensively. The analytical methods are demonstrated in manageable chunks, via case studies, as described above. The final chapter returns to the original questions and aims of the book, armed with insights from the detailed explorations in between, and offers a description of the symbiotic relationship between stylistics and corpus linguistics.

In general, this book should prove to be a very useful resource for researchers and supervisors. The backgrounds to theoretical issues and current debates are presented with clarity and thoroughness. Corpus methods, including widely used tools and processes, which might be new to some stylisticians, are explained with an abundance of data, with handy pointers to online resources. Features which are particularly successful in introducing a new approach to existing stylisticians include the detailed case studies, the step-by-step worked examples and the practice of introducing a concept very briefly in one chapter, before going into detail in subsequent chapters, with useful signposting. The reader can be reassured that they are being introduced to the most central concepts in the area, via the classic studies, and are being exposed to probably the most widely used software and corpora. This introduction has tremendous scope, but also depth, and has staked out a corner of the field of corpus linguistics for stylistics.

References

- Anthony, L. (2018). AntConc (Version 3.5.6) [Computer Software]. Tokyo, Japan: Waseda University. <https://www.laurenceanthony.net/software>
- Davies, Mark. (2004) British National Corpus (from Oxford University Press). <https://www.english-corpora.org/bnc/>
- Freeman, D.C. (1970) 'Linguistics approaches to literature', in Freeman, D.C. (ed.) *Linguistics and Literary Style*, pp.3-17. New York: Holt, Reinhart and Winston.
- Johansson, S., Leech, G.N. and Goodluck, H. (1978) *Manual of Information to Accompany the Lancaster-Oslo/Bergen Corpus of British English, for Use with Digital computers*. Department of English, University of Oslo.
- McIntyre, D. & Walker, B. (2019) *Corpus Stylistics: Theory and Practice*. Edinburgh University Press.
- Rayson, P. (2009) Wmatrix: a web-based corpus processing environment, Computing Department, Lancaster University. <<http://ucrel.lancs.ac.uk/wmatrix/>>
- Scott, M. (2016) *Wordsmith Tools, Version 7*. Stroud: Lexical Analysis Software.